B5

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A01K 67/027, C12Q 1/68

A1

(11) International Publication Number: WO 98/23145

(43) International Publication Date: 4 June 1998 (04.06.98)

(21) International Application Number: PCT/US97/21360

(22) International Filing Date: 25 November 1997 (25.11.97)

(30) Priority Data:
08/756,076 25 November 1996 (25.11.96) US

(71) Applicant: GENE/NETWORKS INCORPORATED [US/US]; Suite 360, 150 North Robertson Boulevard, Beverly Hills, CA 90211 (US).

(72) Inventors: LUSIS, Aldons, J.; 3368 Wade Street, Los Angeles, CA 90066 (US). ANDALIBI, Ali; 10780 Wilkins Avenue, Los Angeles, CA 90024 (US).

(74) Agent: HALLUIN, Albert, P.; Howrey & Simon, 1299 Pennsylvania Avenue, N.W., Box 34, Washington, DC 20004–2402 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: INDEXED LIBRARY OF RECOMBINANT CONGENIC NON–HUMAN ANIMALS

(57) Abstract

Libraries of genetically tagged recombinant congenic non–human animals are described that are useful for identifying the genetic basis of disease. The library shall preferably comprise, *inter alia*, recombinant congenic strains that have been verified as containing substantially discrete/contiguous regions of the donor parent genome in the context of at least one background genome. Additionally described are recombinant congenic non–human animals and strains, as well as methods for generating and detecting the same, that comprises genetically defined and contiguous regions of donor parent genome in the context of at least one background genome.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INDEXED LIBRARY OF RECOMBINANT CONGENIC NON-HUMAN ANIMALS

1.0. Field of the Invention

The present invention relates to indexed libraries of recombinant congenic non-human animals, and methods for producing the same. The animals comprising the library are useful for the generation of offspring with specific genetic traits and phenotypes, as well as the study of the genetic basis of disease.

2.0. Background of the Invention

Standard breeding practices have been developed and used over centuries of trial and error to develop a wide variety of domesticated animal breeds. The genetic similarity of the animals of a given breed results in a population of animals that, as a group, tend to have fairly uniform and predictable behavioral and growth characteristics. Unfortunately, classical "pure breed" animals do not breed true, and a significant proportion of the progeny of "pure breed" crosses often display a wide variety of abnormalities as an adverse consequence of poorly controlled in-breeding.

Unlike "pure-bred" animals, inbred strains comprise animals that have been bred to contain essentially identical genetic content. Consequently, animals within a given inbred strain may generally be expected to breed true. Once inbred animals had been developed, the next phase in animal genetics involved the genetic construction of strains that contained discrete portions of the genetic material from one strain in the context of the genetic background of a second strain. Such strains are typically referred to as "congenic strains".

Congenic animals have widely benefitted scientific research in fields such as immunology. Commercially available congenic mice generally have a relatively minor portion of the genome from the donor strain as compared to the genetic "background" strain. Consequently, even when considered in total, existing congenic strains only represent about a third of the mouse genome. Accordingly, only a

limited percentage of the mouse genome may be studied using the congenic mouse strains that are presently available.

3.0. Summary of the Invention

The present invention is directed to methods for systematically constructing a novel library of non-human recombinant congenic animals. More particularly, two relatively divergent (and preferably highly divergent) inbred strains of animals are identified or constructed, and these strains are then used to construct a library of recombinant congenic animals that collectively transfers a comprehensive set of the donor strain genome into the background strain. Generally, a reciprocal series of crosses shall complete the library by producing a substantially inverse set (i.e., reversing the relative "donor" and "backgrounds" of the first series of crosses) of recombinant genetic animals.

Accordingly, one embodiment of the present invention is a library of non-human recombinant congenic animals comprising animals having genomes containing about 1.5 to about 10 percent donor parent genome. In a preferred embodiment, the presently described libraries shall comprise recombinant congenic strains having discrete and overlapping segments of donor genetic material that collectively constitute a substantially complete representation of the donor parent genome.

Another aspect of the present invention are libraries of "indexed" or "genome tagged" recombinant genetic animals.

Another embodiment of the present invention is the use of the animals in the recombinant congenic library to identify and analyze the genetic basis of disease (or absence thereof). More particularly, an appropriately constructed recombinant congenic library allows for the identification and study of the gene or genes associated with a given phenotype. Accordingly, an additional aspect of the present invention is the generation and use of recombinant congenic animals that comprise regions of the donor genome that are sufficiently specific or defined to allow for fine structure

genetic mapping of phenotypic traits via the generation and characterization of animals that are congenic for specific genetic loci.

Additional embodiments of the claimed invention are methods of screening recombinant genetic libraries by phenotypic screening followed by genetic fine structure mapping of the genetic basis for the selected phenotype.

An additional embodiment of the claimed invention is a recombinant congenic animal that essentially comprises at least a single chromosome (diploid or haploid) corresponding to donor DNA with the remaining chromosomes essentially derived from a background strain (i.e., consomic). Preferably, such an animal/strain shall be homozygous for substantially all of the alleles contained in the donor chromosome (as verified by genome tagging).

Another embodiment of the present invention is a recombinant congenic animal that comprises a substantially contiguous (as opposed to mosaic) portion of between about 1 to about 150 centimorgans of donor genome in the context of a different genetic background.

Additionally, another embodiment of the present invention are inbred recombinant congenic strains that comprise substantially contiguous (as opposed to mosaic) segments (of about 1 to about 150 centimorgans) of donor parent DNA in the context of at least one different genetic background.

Another embodiment of the present invention is the use of the recombinant congenic strains of the library to breed or construct animals, particularly mice, that are useful as human disease models, or for the testing of pharmaceuticals and new chemical entities.

4.0. Detailed Description of the Invention

The presently defined invention allows for a rapid and comprehensive method for associating a given phenotype with the underlying genetic basis for the phenotype. More particularly, the present invention describes methods and

rationales for producing a library of recombinant congenic "genome-tagged" animals. The individual animal strains that collectively constitute the library may each contain either mosaic (i.e., interspersed throughout the genome) or discrete/defined (substantially contiguous) percentages/segments of a "donor" strain genome that have been placed in the genetic context of the "background" strain. For the purposes of the present invention the term "donor strain" shall refer to the strain of animal contributing the minority of the genome of a given recombinant congenic strain, and the terms "background strain" or "recipient strain" shall refer to the strain of animal contributing the majority of the genome of a given recombinant congenic strain.

The presently described recombinant congenic strains are generally constructed by outcrossing two different inbred parental strains followed by a series of back-crosses and an extended series of brother/sister matings to produce recombinant congenic animals that comprise the desired percentages of the donor and background genomes. For example, statistics predict that about 25 percent of genomes in offspring of a first back-cross will be derived from the "donor" genome. Subsequent rounds of back-crossing may be statistically predicted to produce offspring comprising about 12.5, 6.25, 3.125, or 1.5625, etc. percent donor genome. After back-crossing, the genetic composition of the back-crossed offspring is typically characterized by an extended series of genetic screens. Animals bearing the desired genotypes are subject to an extended series of brother/sister matings in order to produce inbred recombinant congenic strains that are homozygous for the desired genotypes or "tagged" genomes.

During back-crossing as well as the construction of recombinant congenic and congenic strains, a variety of screening processes may be used to confirm both the relative percentages of donor/background genetic information, as well as to determine the specific regions of donor and background

DNA contained within each recombinant congenic strain. Methods of screening that may be used to assess the genetic make-up of an animal include, but are not limited to: polymerase chain reaction, genome mismatch scanning, restriction site polymorphism, sequence analysis, DNA hybridization, restriction landmark genome scanning (RLGS), restriction fragment length polymorphisms (RFLP), single-stranded conformation polymorphisms (SSCP), etc.

By identifying the genetic markers present in a given animal, the above screening procedures essentially provide a genetic index for each animal tested. Thus, the fact that the individual animals/strains in the library have been comprehensively genome-tagged is a novel feature of the presently described libraries.

For the purposes of the present disclosure, the terms "indexing" or "genome-tagged" refer to the processes, during screening, by which genetic markers are tracked, catalogued, or otherwise inserted in order to identify the portions of the donor genome present in a given animal or strain. Once animals with the desired genetic make-up are identified during the genetic screening process, the desired genotypes are rendered homozygous by brother/sister matings, thereby constructing the appropriate recombinant congenic strains. Subsequent rounds of the genome screening/tagging process are used to verify the desired genotypes and to ensure that essentially all of the donor genome is collectively represented in the library of recombinant congenic animals. For the purposes of the present invention, a "substantially genome tagged animal or strain" has effectively been screened to an extent that a genetic map has been generated that accounts for at least one genetic marker of interest in every contiguous 1 to about 100 centimorgan region of the genome, preferably for at least one genetic marker in every 2 to about 75 centimorgans of the genome, more preferably about 5 to about 65 centimorgans, and specifically at least one genetic marker in about every 50 centimorgans of the genome. It shall be understood that a given animal or strain may not

need to be comprehensively screened/tagged over regions in which both parents are known to homozygously encode the same genetic content, and thus a given animal/strain may be effectively "substantially genome tagged" even though only a portion of that animal or strain's genome has actually been genetically mapped. A given animal or strain shall be considered "comprehensively genome mapped" where one can account for at least 1 genetic marker in about every 1 to 10 centimorgans of the genome, preferably about 2 to about 8 centimorgans, and specifically about 4 to about 6 centimorgans of the genome (excluding telomeres).

In order to facilitate the screening processes (e.g., genome mismatch scanning, etc.) used to genome tag the recombinant congenic animals, it is advantageous to use parental strains that are genetically divergent to increase the chances that polymorphisms may be detected that allow one to genetically distinguish the parental strains. For the purposes of the present invention the term "genetically divergent" shall mean that the animals of interest shall generally contain an average of at least about 0.25 percent genetic mismatch, preferably at least about 0.5 percent, and more preferably at least about 0.75 percent. Typically, "highly divergent" strains shall comprise genetic backgrounds that differ by at least about 1.0 percent (i.e., the DNA sequences of the strains differ by an average of at least about 1 out of every 100 nucleotides).

Typically, the library shall comprise, *inter alia*, recombinant congenic animals having about twelve (12) percent of their genomes derived from the donor strain. Where such animals are used to construct a library, the number of the animal strains necessary to generate a set of recombinant congenic strains that collectively represent an essentially complete copy of the "donor strain" genome shall vary from about nine strains to several dozen or more (i.e., hundreds) strains depending upon the percentage of donor genome in each strain and the extent of genetic overlap that one may wish to engineer into the recombinant congenic strains that

collectively constitute the library. For the purposes of the present invention, the term "substantially complete" shall mean that a statistical analysis may show that there is at least about a 95 percent probability that at least about 95 percent of the donor genome is collectively represented in a given library.

Preferably, the library of recombinant congenic animals will comprise strains that contain between about 0.5 percent and about 12 percent of the donor strain genome, more preferably, between about 1 percent and about 8 percent, specifically between about 1.5 percent and about 6.2 percent, and more specifically about 2 percent to about 4 percent of the donor strain genome, or any mixture or combination of the above. Alternatively, the library may comprise recombinant congenic strains having between about 5 to about 100 centimorgans, typically about 10 to about 80 centimorgans, preferably about 15 to about 75 centimorgans, more preferably about 25 to about 60 centimorgans of donor genome, and specifically about 50 centimorgans of donor genome. Given that, for example, a typical mouse genome has approximately 1,500 centimorgans of genetic material, 50 centimorgans corresponds to approximately 3 percent of the total mouse genome.

Depending upon the percentage of the donor genetic material transferred into the animals in the library, a corresponding number of overlapping strains are necessarily required to form a library of recombinant congenic animals that collectively represent a substantially comprehensive portion of the donor strain genome. Accordingly, the libraries of the present invention may comprise up to many hundreds of recombinant congenic animals strains. It should also be understood that the above number may effectively double when a complementary (or "mirror image" library is constructed using recombinant congenic animals bearing substantially inverse proportions of genetic material from the "donor" and "recipient" strains. For the purposes of the present invention, the term "substantially inverse" shall

mean that similar methods have been used to construct a series of recombinant congenic strains that reverse the "donor" and "recipient" genetic components.

The recombinant congenic strains of the presently described library are typically subject to an extensive series of genetic screens. The result of this screening process is that each strain in the library will have been analyzed and characterized over a region spanning essentially the entire genome. Accordingly, each strain in the library will have been comprehensively evaluated or "genome tagged". Consequently, the results of the genetic analysis (or genome tagging) may be compiled and stored in a database that, *inter alia*, shall also serve as a genetic "index" to the recombinant congenic strains in the library.

An additional feature of the genomic screening process is that recombinant congenic strains shall be identified that comprise defined or discrete regions of the donor genome. For the purposes of the present invention, a discrete region of genome shall mean that a substantially contiguous series of genetic markers shall have been derived from one of the parent genomes. Accordingly, recombinant congenic strains that comprise defined or discrete segments shall generally comprise at least about 60 percent, preferably about 75 percent, more preferably at least about 90 percent, specifically at least about 95 percent of the net donor parent genome component (of the total genome) in one to about three regions of substantially contiguous genetic information. Accordingly, a particularly preferred embodiment of the present invention is a library where at least one subset of the strains in the library have substantially contiguous regions of donor genome, and where the substantially contiguous regions of donor genome contained in these strains overlap to an extent that the strains collectively form a substantially complete representation of the donor genome in the context of at least one background genome. For the purposes of the present invention, the term "substantially contiguous" shall mean

that at least about 90 percent of the genetic markers screened within the region correspond to a single parent strain genome. The term "contiguous" (region of genome) shall refer to the situation where all of the genetic markers screened within a given region correspond to a single parent strain genome.

Alternatively, recombinant congenic strains used to generate the presently described library may also comprise genetic mosaics (genomes that intersperse a multiplicity of donor sequences within the background genomes).

In any event, the presently described library of recombinant congenic animals shall allow for the rapid and facile identification of the genetic features that are associated with a given phenotype, or multigenic or polygenic disorder.

Although mice are specifically exemplified herein, the present invention is broadly deemed to have far reaching applications. Virtually any species of animal that is capable of being inbred to an extent that it is essentially homozygous at all loci may serve as the basis for a library of recombinant congenic animals. Accordingly, in addition to use in studying the genetic contribution to disease and disease susceptibility, fields such as animal husbandry, wild life management, and the breeding of domesticated animals as pets may be significantly advanced by the presently disclosed invention. For example, a library of recombinant congenic swine could be used (in conjunction with a database on the genetic tendencies of the offspring of particular crosses) to generate animals with predictable and enhanced body and growth characteristics (i.e., percent body fat, rapid growth, resistance to specific diseases, enhanced skin, fur, or coats, etc.). Since the recombinant congenic strains of the present invention breed true, offspring generated using the library strains may rapidly be converted into strains/breeds that breed true. Accordingly, the present invention effectively provides a method for developing a wide variety of recombinant congenic or inbred strains of animals bearing

desirable and predictable genetic tendencies, that also remain capable of predictably and reliably reproducing offspring sharing the desired qualities.

In view of the above, animals that may be used to construct the presently described recombinant congenic strains may include, but are not limited to, pigs, rats, rabbits, cattle, goats, fish, and birds (particularly turkeys and chickens), as well as animal species, particularly mammalian species, that are known in the art. Additionally, bovine, ovine, and porcine species, other members of the rodent family, e.g. rat, as well as rabbit and guinea pig, and non-human primates, such as chimpanzee, may be used to practice the present invention. Particularly preferred animals are rats, rabbits, guinea pigs, and most preferably mice.

Given that recombinant congenic mice are specifically exemplified herein, an additional embodiment of the present invention is a recombinant congenic mouse, or inbred strain thereof, that generally comprises between 1 to about 150 centimorgans, and preferably comprises about 10 to about 80 centimorgans of the donor genome as a substantially contiguous region (or regions where multiple chromosomes are involved) of donor genome. Preferably, recombinant congenic mice, or strains thereof, bearing the indicated substantially contiguous regions of donor genome will be the offspring arising after at least about 3 rounds, preferably at least about six rounds, and more preferably at least about a dozen rounds of brother-sister matings. In a particularly preferred embodiment, the presently described recombinant congenic mice will have been screened to verify the substantially contiguous nature of the donor parent genetic component, before, during, and after (or all or any combination thereof) the requisite series of brother-sister matings.

A particularly preferred embodiment of the present invention includes the use of the presently described library of recombinant congenic strains to study the genetic basis of

disease. Present techniques of classical genetic analysis are generally limited to non-epistatic genetic traits. Accordingly, classical genetic techniques are only of limited value in instances where a given disease or phenotype is dependent on the interaction of multiple genetic loci. The presently described libraries of recombinant congenic animals provide the tools necessary for a systematic analysis of the genetic component of a variety of animal diseases or developmental and behavioral characteristics. Additionally, because of the substantial amount of donor genetic material present in the various strains in the library, the presently disclosed strains are particularly well suited for the study of multigenic and polygenic genetic disorders or tendencies.

In particular, the novel genetic content of the strains in the library should correlate with the observation of equally novel phenotypic tendencies (with or without measures to promote a given phenotype). Additionally, it has been shown that a given animal strain may serve as a useful and informative human disease model when subject to the proper environmental stimuli or stress. Accordingly, an additional embodiment of the present invention includes the employment of a systematic method of identifying and subjecting an animal to a series of environmental stimuli or stresses in order to determine whether the animal, when subject to a specific stress or environment, develops symptoms or signs that are similar to, or associated with, a known disease in other animals including humans. Examples of such stresses or environmental stimuli may include, but are not limited to: specialized (e.g., high-protein or high fat, etc.) diets, exposure to toxic agents or chemicals, lethal or sub-lethal irradiation, drug treatments, hyperthermia or hypothermia, viral infection, bacterial infection, exercise tolerance and training, response to inhalation of various agents (e.g., smoke, asbestos, airborne chemical toxins, etc.), behavioral responses (e.g., maze training), circadian rhythm changes, physical shock, surgeries, behavioral stimuli, sound, natural and synthetic pharmaceuticals or chemical agents, exposure to

foreign DNA, immune challenge, immunocompromising agents or procedures, etc.

Typically, more extensive and detailed phenotyping procedures are more likely to identify a relationship between a disease symptom and the responsible gene or genes. One of the most valuable but labor-intensive of the phenotypic measurements is that of temporal, environmental and tissue-specific patterns of gene expression in normal and diseased tissues. Combining such gene expression data with other phenotypic data, as well as the genotypic data from the library, provides a comprehensive approach to gene discovery.

The selection of the proper challenge/stimulus may be done randomly, or may be consistent with previously identified disease associations (e.g., ulcers/*helicobacter* infection, Crohn's disease/*bacteriodes or M. paratuberculosis* infection, etc.).

An additional feature of subjecting the recombinant congenic animals of the library to environmental or dietary stress or challenge is that genotypes are identified that correlate with an enhanced ability to avoid or substantially mitigate the adverse consequences normally associated with such stresses. Accordingly, another embodiment of the present invention relates to the identification and characterization of the genetic basis for an enhanced resistance to normally adverse stimuli (e.g., high fat diets, sleep deprivation, noxious inhalants, etc.). Additionally, the presently described methods and libraries are equally well suited for identifying the molecular basis for genetically determined advantages such as prolonged life-span, low cholesterol, low blood pressure, exercise tolerance, optimal respiratory quotient, low percent body fat, reduced bone absorption, increased memory, attenuation of sensitivity or prevention of inflammatory or autoimmune disorders, etc.

An additional feature of the present invention is that it also provides means of identifying genes that do not actually cause a disorder, but modify or mediate the severity

of the phenotype or disorder. Examples of such situations include, but are not limited to, genes that alter age at disease onset, or genes that affect the progression and outcome of neoplasms (e.g., metastasis).

Once a relevant disease pathway has been identified, other recombinant congenic strains in the presently described libraries may be similarly stressed to ascertain whether the disease symptoms correlate with a particular region (or regions) of the genome. Where such a correlation is seen, the overlapping and reductive nature of the recombinant congenic strains in the library allow for the rapid fine-structure genetic analysis of the disease system. Alternatively or additionally, after a given genomic region has been associated with a given phenotype, fine-structure genetic mapping of the phenotype may be conducted by using the relevant recombinant congenic animals to construct true congenic animals (by an additional series of back-crosses) in order to narrow the precise genetic interval containing the gene of interest. The resulting congenics animals may subsequently be used to more specifically identify and isolate the loci that are associated with the phenotype of interest.

Given the sensitivity of the genetic tools available through the library, the eventual identification and isolation of the specific gene or genes that are associated with the disease symptoms may be accomplished more rapidly and efficiently. For example, recombinant congenic strains necessarily delimit the area of genetic analysis. Thus, once a phenotypic trait has been associated with a given recombinant congenic strain, one significantly reduces the time and expense necessary to positionally clone a given gene via established methodologies (e.g., chromosome walking, YAC construction, exon trapping, polynucleotide or cDNA hybridization or sequence analysis, etc.).

Once a putative gene or other region of the chromosome has been isolated, the gene may subsequently be analyzed using established molecular biology techniques (i.e.,

cloning, DNA sequencing, polymerase chain reaction, etc., see generally Sambrook *et al.* (1989) Molecular Cloning Vols. I-III, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, and Current Protocols in Molecular Biology (1989) John Wiley & Sons, all Vols. and periodic updates thereof, herein incorporated by reference). Additionally, once such sequences have been isolated and characterized, they may be used as heterologous probes to rapidly and efficiently identify and isolate related sequences in other organisms and animals.

The presently disclosed methods, libraries, and databases are useful for the diagnosis, study, and treatment of a wide variety of diseases and disorders including, but not limited to: autoimmune disease, arthritis (e.g. rheumatoid and osteo), atherosclerosis, obesity, osteoporosis, microbial or viral infection, inflammatory bowel diseases, diabetes (insulin dependant or non-insulin dependant), multiple sclerosis, cancer, Crohn's disease and ulcerative colitis, asthma, hypertension, dementia, Alzheimer's disease, stroke, hepatitis, chronic liver disease, peptic ulcers, gastritis, neurodegenerative disorders, neurological or psychiatric disorders, systemic lupus erythematosus, scleroderma, post-angioplasty vascular inflammatory response, etc.

Cancers that may be diagnosed or treated using the presently described methods include, but are not limited to: Breast: sarcoma and carcinoma; Cardiac: sarcoma (angiosarcoma, fibrosarcoma, rhabdomyosarcoma, liposarcoma), myxoma, rhabdomyoma, fibroma, lipoma and teratoma; Lung: bronchogenic carcinoma (squamous cell, undifferentiated small cell, undifferentiated large cell, adenocarcinoma), alveolar (bronchiolar) carcinoma, bronchial adenoma, sarcoma, lymphoma, chondromatous hamartoma, mesothelioma; Gastrointestinal: esophagus (squamous cell carcinoma, adenocarcinoma, leiomyosarcoma, lymphoma), stomach (carcinoma, lymphoma, leiomyosarcoma), pancreas (ductal adenocarcinoma, insulinoma, glucagonoma, gastrinoma,

carcinoid tumors, vipoma), small bowel (adenocarcinoma, lymphoma, carcinoid tumors, Karposi's sarcoma, leiomyoma, hemangioma, lipoma, neurofibroma, fibroma), large bowel (adenocarcinoma, tubular adenoma, villous adenoma, hamartoma, leiomyoma); Genitourinary tract: kidney (adenocarcinoma, Wilm's tumor [nephroblastoma], lymphoma, leukemia), bladder and urethra (squamous cell carcinoma, transitional cell carcinoma, adenocarcinoma), prostate (adenocarcinoma, sarcoma), testis (seminoma, teratoma, embryonal carcinoma, teratocarcinoma, choriocarcinoma, sarcoma, interstitial cell carcinoma, fibroma, fibroadenoma, adenomatoid tumors, lipoma); Liver: hepatoma (hepatocellular carcinoma), cholangiocarcinoma, hepatoblastom, angiosarcoma, hepatocellular adenoma, hemangioma; Bone: osteogenic sarcoma (osteosarcoma), fibrosarcoma, malignant fibrous histiocytoma, chondrosarcoma, Ewing's sarcoma, malignant lymphoma (reticulum cell sarcoma), multiple myeloma, malignant giant cell tumor, chordoma, osteochronfroma (osteocartilaginous exostoses), benign chondroma, chondroblastoma, chondromyxofibroma, osteoid osteoma and giant cell tumors; Nervous system: skull (osteoma, hemangioma, granuloma, xanthoma, osteitis deformans), meninges (meningioma, meningiosarcoma, gliomatosis), brain (astrocytoma, medulloblastoma, glioma, ependymoma, germinoma [pinealoma], glioblastoma multiforme, oligodendroglioma, schwannoma, retinoblastoma, congenital tumors), spinal cord (neurofibroma, meningioma, glioma, sarcoma); Gynecological: uterus (endometrial carcinoma), cervix (cervical carcinoma, pre-tumor cervical dysplasia), ovaries (ovarian carcinoma [serous cystadenocarcinoma, mucinous cystadenocarcinoma, endometrioid tumors, celioblastoma, clear cell carcinoma, unclassified carcinoma], granulosa-thecal cell tumors, Sertoli-Leydig cell tumors, dysgerminoma, malignant teratoma), vulva (squamous cell carcinoma, intraepithelial carcinoma, adenocarcinoma, fibrosarcoma, melanoma), vagina (clear cell carcinoma, squamous cell carcinoma, botryoid sarcoma [embryonal rhabdomyosarcoma], fallopian tubes

(carcinoma); Hematologic: blood (myeloid leukemia [acute and chronic], acute lymphoblastic leukemia, chronic lymphocytic leukemia, myeloproliferative diseases, multiple myeloma, myelodysplastic syndrome), Hodgkin's disease, non-Hodgkin's lymphoma [malignant lymphoma]; Skin: malignant melanoma, basal cell carcinoma, squamous cell carcinoma, Karposi's sarcoma, moles, dysplastic nevi, lipoma, angioma, dermatofibroma, keloids, psoriasis; and Adrenal glands: neuroblastoma.

Given that the disclosed libraries allow for the rapid identification of the genetic basis for a disease, an additional aspect of the present invention is the identification and cloning of human homologues to genes identified using the presently disclosed libraries of recombinant congenic strains. Ideally, the human homologue of a given gene will hybridize with the heterologous probe under stringent conditions. For the purposes of the present invention "stringent conditions" generally refers to hybridization conditions that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C.; or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M Sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 g/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC and 0.1% SDS. The above examples of hybridization conditions are merely provided for purposes of exemplification and not limitation. A more thorough treatise of the such routine molecular biology techniques may be found in Sambrook *et al.*, Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Vols. 1-3: (1989), and periodic updates thereof, herein incorporated by reference.

Once a genetic locus of interest has been identified, an additional embodiment of the present invention is the testing of the observed genetic correlation via the production of transgenic animals that have been engineered to be deficient in (or to over express) one or more alleles of the gene of interest. Typically, in the case of knockout transgenics, such transgenic animals are produced by using disruptive recombination in embryonic stem cells that are subsequently incorporated into blastocysts. Alternatively, in the case of expression-enhanced transgenics, DNA is injected into the male pronucleus of a fertilized egg and implanted into pseudo pregnant females. Such gene-modified animal cells and transgenic animals may be prepared using any of several techniques that are well established in the art. In particular, techniques may be used that are conceptually similar to those taught in U.S. Patent No. 5,464,764 issued to Capecchi and Thomas on November 7, 1995, herein incorporated by reference.

Other embodiments of the present invention include cell lines derived from the recombinant congenic animals of the presently described libraries. Once a genotype has been associated with a disease, disorder, or advantage, the precise molecular basis of the disorder may be best analyzed at the cellular level. Accordingly, cell lines derived from tissues manifesting the adverse or advantageous genetic tendencies may be used to more particularly study the chemical basis for a given disease or advantage.

In many instances, different cellular genotypes may exhibit markedly different patterns of gene expression, and different tissues within a genotype may also be differentially effected. Thus, an additional embodiment of the presently disclosed invention is the development of expression profiles for recombinant congenic cells and recombinant congenic animals with varying genetic content. Additionally, expression profiles may be determined for different tissues having the same genotype, or for similar tissues having the same genotype which were obtained from

animals that may or may not display a given phenotypic trait. For the purposes of the present invention, the term "expression profile" shall refer to a sequence database that provides quantitative and qualitative information regarding the genes expressed by a given animal, or cells obtained or derived therefrom. Generally, such profiles may be generated by isolating mRNA from ("affected" and "control") animal tissue or cells, reverse transcribing the mRNA, constructing a cDNA library, sequencing the clones in the cDNA library, and gathering and organizing, and analyzing the DNA sequence data. A comparison of the expression profiles for tissues or cells from animals with different genotypes provides a direct measure of the cellular consequences of the genetic variance.

Alternatively, an expression profile may be generated by generating the appropriate cDNA library, and screening the cDNA library with differentially labeled RNA from normal and affected animal tissue. The clones whose expression are altered (in comparison to normal) are then identified, sequenced and further analyzed.

Additionally, chemical mutagenesis may be used to effectively generate new alleles at a given genetic locus, and the expression profiles of mutagenized animals or cells may be screened as discussed above.

A still further aspect of the invention relates to constructing subtraction-libraries generated by using transcripts from nonstimulated or nonresponder samples to subtract transcripts that are not responsible for the trait of interest. Subtraction-libraries may be constructed from a source of interest using techniques which are standard in the art (Molecular Cloning: A Laboratory Manual, 2d ed., Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989)). For example, a subtraction-library may be constructed that will identify those cDNAs which are "new" in a recombinant congenic strain exhibiting the trait of interest by subtracting out the cDNAs that are present in control cells. Such cDNAs constitute ideal probes for identification and

localization of the gene or genes responsible for the phenotypic trait of interest.

Alternatively, standard subtraction analysis may be combined with gene expression profiling of the genes that are turned on, turned off, or modulated in disease states. After a modulated gene has been identified and associated with a trait of interest, further analysis may reveal a coherent and detailed picture of the sequence of events that occur during the disease process. This approach provides a powerful complement to standard positional cloning methods.

High-throughput and extremely sensitive assay systems allow for the construction of detailed, tissue-specific gene expression profiles in animals such as humans and mice. In particular, a technology platform based on automated differential gene expression analysis is exemplified. Such platforms have previously involved hybridizing cDNAs derived from "normal" and "abnormal" tissues to arrays of cDNAs from a "target" cDNA library. The data obtained from such a platform may then be used in further gene discovery efforts.

The following examples are provided solely to illustrate the subject invention. Given the level of skill in the art, one may be expected to modify any of the above or following disclosure to produce insubstantial differences from the specifically described features of the present invention. As such, the following examples are provided by way of illustration and are not included for the purpose of limiting the invention in any way whatsoever.

5.0. EXAMPLES

5.1. Library construction

5.1.1. Animal Strains

In order to highlight the desired phenotypic effects and facilitate the genetic screening process, it is generally desirable to construct a genome tagged recombinant congenic library using relatively divergent animal strains.

Where mice are used to construct a library, examples of such divergent strains may be found within the *Mus musculus*

group. For example, *Mus musculus domesticus* (Domesticus) and *Mus musculus castaneus* (Castaneus) are highly divergent mouse subspecies. In order to facilitate phenotypic screening, it is also generally desirable to utilize strains that have been previously characterized as possessing known phenotypic traits. Additionally, it is desirable to select donor and background strains on the basis of having many phenotypic differences between them (e.g., atherosclerosis susceptibility, obesity, bone density, etc.).

Apart from its genetic distance, Castaneus is also phenotypically distinct from many domesticus strains. For example, Castaneus has a much higher bone density than the Domesticus strain C57BL/6, which would make a recombinant congenic animal from the two strains an ideal candidate model for a genetic study of osteoporosis and bone density differences.

The following is a partial listing of some of the phenotypic characteristics of several Domesticus strains that may be used to construct the described libraries:

a. <u>C57BL/6</u>

Susceptible to the development of atheromatous lesions on wall of aorta after 20 weeks on a high-fat diet (Thompson, 1968; Roberts and Thompson, Paigen, 1995). Develop fatty streak-like lesions in the valve sinus region of the ascending aorta after 10-20 weeks on a diet enriched in saturated fat and cholesterol. After a further 15 weeks, fibro-fatty lesions are found with many of the characteristics of human atheromatous plaques (Stewart-Phillips and Lough, 1991, Atherosclerosis *90*:211-218). Develops non-insulin-dependent diabetes mellitus and hypertension when fed a high fat-high simple carbohydrate diet, whereas A/J mice do not (Mills *et al.*, 1993, Am. Jour. Phys. *264*:R73-8). Low plasma cholesterol at 12 and 24 weeks. Low plasma triglyceride levels and low plasma cholesterol (Jiao *et al.*, 1990, Metabolism *39*:155-160). Low serum ceruloplasmin levels in males but intermediate in females

(Meier and MacPike, 1968, Proc. Soc. Exp. Biol. Med. *128*:1185-1190). High blood sugar (Nishimura, 1969, Exp. Animals (Japan) *18*:147-157). High liver tyrosine aminotransferase in fasted mice but low in C57BL/6-ob (Blake, 1970, Int. J. Biochem. *1*:361-370). Low calcium uptake by the heart (Mokler and Iturrian, 1973, Proc. Soc. Exp. Biol. Med. *142*:919-923). High aldehyde dehydrogenase and alcohol dehydrogenase activity compared with DBA/2 (Sheppard *et al.*, 1968, Biochem. Genet. *2*:205-212). Low basal levels of kidney catalase, superoxide dismutase, and renal glutathione reductase (Misra *et al* 1991, Toxicology Letters *58*:121-33). High alcohol preference (Fuller, 1964, J. Comp. Physiol. Psychol. 57: 85-88; Rodgers, 1966, Psychosomat. Med. *28*:498-513). Low "emotionality", high open-field exploration (Thompson, 1953, Can. J. Psychol. 7:145-155). High preference for sweet tasting substances (Lush, 1988, Genetic. Res. Camb. *53*:95-99). Low susceptibility to audiogenic seizures (Deckard *et al.*, 1976, Develop. Psychobiol. *24*:9-17).

b. A/J

Can be made obese by a suitable diet (Fenton and Dowling, 1953, J. Nutrition *49*:319-331). Does not develop non-insulin-dependent diabetes mellitus and hypertension when fed a high fat-high simple carbohydrate diet, whereas C57BL/6 mice do (Mills *et al.*, 1993). High incidence of amyloidosis (Russell and Meier, 1966, In Biology of the Laboratory Mouse, 2d, (Green, ed.), McGraw-Hill, New York, pp. 571-587). Low metabolic rate (Storer, 1967, Exp. Gerontol *2*:173-182). Low serum ceruloplasmin in males but intermediate in females (Meier and MacPike, 1968). Responds by higher growth rate on high fat diets (Fenton and Carr, 1951, J. Nutrition *45*:225-233). High percent carcass lipid on a high-fat diet (West *et al.*, 1992). Develops autoimmune phenomena, immunological deficits with aging and autoimmunity following neonatal thymectomy (Yunis *et al.*, 1972, The thymus, autoimmunity and

the involution of the lymphoid system, (ed. Sigel, M.M. and Good, R.A.), Thomas, Springfield, Ill., pp. 62-119).

c. DBA/2J

Susceptible to audiogenic seizures (Fuller and Sjursen, 1967, J. Hered. *58*:135-140). Spontaneous calcified heart lesions progress with age - 90% of individuals affected by 1 year (Rings and Wagner, 1971, Lab. Animal Sci. *22*:344-352). Incidence of calcareous heart lesions high among some related strains (Di Paola *et al.*, 1964, Proc. Soc. Exp. Biol. Med. *115*:496-497). Other lesions include malignant lymphoma and degenerative processes in the myocardium, skeletal muscle, subcutaneous adipose tissue, cornea and blood vessels. Lesions partly depend on diet (Hare and Stewart, 1956, J. Nat. Cancer Inst. *16*:889-911). High metabolic rate (Storer, 1967, Exp. Gerontol *2*:173-182). High concentration of epinephrine and norepinephrine in adrenals (Ciaranello *et al.*, 1972, Life Sci. *11(1)*:565-572). Low liver tyrosine aminotransferase activity in fasted mice (Blake, 1970). High calcium uptake by the heart (Mokler and Iturrian, 1973). Low aldehyde and alcohol dehydrogenase activity compared with C57BL/6 (Sheppard *et al.*, 1968).

d. C3H/HEJ

Can be made obese by a suitable diet (Fenton and Dowling, 1953). Low blood pressure (Mullink *et al.*, 1975, Personal Communication). High plasma cholesterol and triglycerides. High erythrocyte catalase (Hoffman and Rechcigl, 1971, Enzyme *12*:219-2225). Low metabolic rate (Pennycuik, 1967, Aust. J. Biol. Med. Sci. *45*:331-346). Low adrenal corticosteroid production (Nandi *et al.*, 1967, Endocrinology *80*:576-582). Low basal level of renal glutathione S-transferase but high basal level of renal glutathione reductase (Misra *et al.* 1991). Has poor immune response to endotoxic lipopolysaccharide due to a B-cell deficit (Rosenstreich and Glode, 1975, J. Immunol. *115*:777-780; Coutinho, 1976, Scand. J. Immunol. *5*:129-140). High

liver tyrosine aminotransferase level in fasted mice (Blake, 1970). Good short-term but poor long-term memory in contrast with DBA/2 (Bovet et al., 1969, Psychopharmacologia *10*:1-5).

5.1.2. Methods Used to Produce Recombinant Congenic Animals

Standard genetic outcrosses are used to initially cross the parental strains to produce offspring that are comprised of 50 percent of each parental genome. Back-crossing the offspring to the relevant parent results in offspring with genomes that are 25 percent "donor"-derived. Subsequent rounds of back-crossing are used to produce successive rounds of offspring that reduce the amount of donor-derived genetic information by one-half per each back-cross. Once offspring with the appropriate genetic content are identified, an extended series of brother-sister matings are used to "fix" the genotype by generating recombinant congenic strains that are homozygous for the desired genotype. The resulting strains are subsequently rescreened to verify the qualitative and quantitative features of their respective genotypes.

5.1.3. Methods Used to Analyze Recombinant Congenic Animals

Previously, back-crosses and intercrosses have been used to locate loci affecting a wide variety of phenotypes. However, fine genetic mapping of the loci of interest requires the generation of hundreds-to-thousands of progeny, which are each genotyped and phenotyped. Using molecular biology techniques to screen for the presence and organization of the relevant genetic tags markedly simplifies and expedites the screening procedure, and allows for the use of methods of data generation and storage that are readily adaptable to automation. A particularly preferred method of genotyping that is well suited for automation is microsatellite mapping.

Microsatellites are highly polymorphic genetic elements that are present at high density throughout all mammalian

genomes studied. Also known as simple sequence repeats (or SSR), microsatellites consist of mono-, di-, tri- or tetrameric sequences repeated multiple times in a tandem array. Allelic variation arises from differences in the number of repeats present in a given tandem array, which results in variation in the total length of the microsatellite. Typically, the length of a satellite is measured by PCR amplification of the microsatellite followed by gel electrophoresis of the product to determine the size. Unique nucleotide sequences on each side of the repeat are chosen as primers for the PCR amplification reaction. The unique flanking sequences also serve to locate, or map, a particular microsatellite on the genome studied. In this way a microsatellite marker becomes a tag for that position on the genome. A large number of simple sequence length polymorphisms or microsatellite markers have been mapped throughout the mouse and human genomes (Dietrich *et al.*, 1996, Nature *380*:149-152; Dib *et al.*, 1996, Nature *380*:152-154 both of which are herein incorporated by reference), thereby generating a linkage map of the markers for each chromosome in the genome.

Thus, an animal may be genotyped by screening and tagging the parental alleles that are present in the animal's genome. Typically, a genome is initially tagged using several hundred of the polymorphic microsatellite markers that effectively cover the entire genome. This method has already been used to genotype several different strains of mice (Whitehead/MIT Genome Center database).

In constructing the libraries, the genotype of an offspring is determined as described above and compared to the genotype of the two parental strains using about 100-400 microsatellite markers, preferably about 350, that are dispersed throughout the genome. By this means, offspring are identified that contain the desired percentage of the donor parent genome, and the desired region or regions of the donor parent genome. These offspring are subsequently sibling-mated to construct a recombinant congenic strain that

is homozygous for the desired genotype. The donor region in each strain is finely mapped using either existing or newly identified markers. Such mapping is used to define the boundaries of the chromosomal region from the donor parent genome, to further confirm that the donor region is contiguous, and to create a fine genetic map of the donor genome segment.

It must also be noted that at any step during the above process, the offspring may be phenotyped for any trait of interest. Offspring may be chosen for further backcrossing and sibling-mating if they display a desired phenotype.

Alternatively, the recombinant congenic strains are constructed, screened, and used to assemble a genome tagged library comprising strains having overlapping regions of donor genome (which may or may not be contiguous or substantially contiguous), in the context of at least one background genome. The animals in the library may then be used in phenotypic screening studies.

5.1.4. Quantitative Trait Locus Mapping

Quantitative trait locus mapping (QTL) is a method that has been successfully used for the analysis of many traits, including the identification of modifying genes. This method is desirable because of the availability of numerous polymorphic markers spanning the entire mouse genome. The approach involves phenotyping backcross (or F2) progeny for the trait of interest as well as genotyping with genetic markers throughout the genome. Typically, nucleic acid samples obtained from the subject animals are genetically typed and analyzed for association of parental alleles with phenotypes of interest.

5.2. Methods of Using the Library to Identify the Genetic Basis for Disease, or the Genetic Basis for a Phenotypic Advantage

5.2.1. Methods of Screening for Phenotypic Effects

Typically, a wide variety of well established animal handling and feeding methods are used to screen recombinant congenic libraries for specific phenotypic effects. Generally, the animals in the library are observed in order to identify any inherent phenotypic traits. Since the animals in the library have been tagged and genetically catalogued during screening, virtually any observed phenotype may be rapidly correlated with a given genotype or particular region of the donor or background genome. Additionally, a representative group of animals from the library may be subject to an appropriate stimuli prior to screening for animals that exhibit distinctive phenotypic traits.

In addition to traditional stimuli such as controlling diet and the environment, more aggressive techniques may be used such as 1,2-dimethylhydrazine injection to study the genetics of susceptibility to formation of aberrant crypt foci and colon adenomas (Moen *et al.* 1996, Can. Res., *56(10)*:2382-6).

As another example, Mori *et al.*, 1995, Genomics *25(3)*:609-14 used recombinant congenic (CcS/Dem) strains derived from inbred mouse strains BALB/cHeA (susceptible) and STS/A (resistant) to study the genetics of susceptibility to radiation-induced apoptosis of thymocytes after whole-body irradiation.

5.2.1. Methods of Generating Expression Profiles

After phenotypic screening has identified a recombinant congenic strain of interest is identified in the library, one can further analyze the regulatory consequences of the genotype by expression profile mapping. Typically, expression profiles are generated by obtaining mRNA samples from the relevant tissues in an appropriately stressed animal, generating cDNA, and using DNA sequence analysis to

quantitatively and qualitatively assess gene expression in the affected tissues. Such expressed sequence tag (EST) expression profiles are useful as a comparative tool versus expression profiles obtained from unstimulated or unaffected tissues or strains.

a) Methods of Using Expression Profiles

The sequence data comprising the expression profile may be used to generate an electronic relational database that may be organized by cataloging the quantity of clones corresponding to a given gene as well as the variety of genes represented in the expression profile. Database programs that are particularly deemed useful for the above analysis may be adapted from commercially available software such as that available from Oracle or Informix, for example.

Of particular interest is the use of sequence, or sequence data, obtained from an non-stimulated, or nonresponder "control" animal to qualitatively or quantitatively "subtract" sequences that are not uniquely represented or over-represented in the strain exhibiting the desired trait or phenotype. Such substraction analysis will drastically reduce the fine structure genetic analysis of the phenotype at issue.

EQUIVALENTS

The foregoing specification is considered to be sufficient to enable one skilled in the art to broadly practice the invention. Indeed, various modifications, variations, and combinations of the above-described methods for carrying out the invention which are obvious to those skilled in the fields of genetics, microbiology, biochemistry, organic chemistry, medicine or related fields are intended to be within the scope of the following claims. All patents, patent applications, and publications cited are herein incorporated by reference.

What is claimed is:

1. A library of non-human recombinant congenic animals, comprising: substantially genome tagged animals having genomes containing about 10 to about 80 centimorgans of donor parent genome.

2. The library of claim 1 wherein said donor parent genome is substantially contiguous.

3. The library of claim 1 wherein said animals are mammalian.

4. The library of claim 3 wherein said mammalian animals are rodents.

5. The library of claim 4 wherein said rodents are a mice.

6. The library of claim 5 wherein said mice collectively represent a substantially complete representation of the donor parent genome.

7. A non-human recombinant congenic animal having a genome comprising about 1 to about 5 percent donor parent genome as a substantially contiguous region.

8. The recombinant congenic animal of claim 7 wherein said animal is a mammal.

9. The recombinant congenic animal of claim 8 wherein said mammal is a rodent.

10. The recombinant congenic animal of claim 9 wherein said rodent is a mouse.

11. An inbred strain of the recombinant congenic mouse of claim 10.

INTERNATIONAL SEARCH REPORT

Inte tional Application No
PCT/US 97/21360

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 A01K67/027 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 A01K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	RAMSAY L.D. ET AL.: "The construction of a substitution library of recombinant backcross lines in Brassica oleracea for the precision mapping of quantitative trait loci." GENOMICS, vol. 39, no. 3, 1996, pages 558-567, XP002058279 see abstract ---	1-11
Y	ANDALIBI A. ET AL.: "Mapping of multiple mouse loci related to the farnesyl pyrophosphate synthetase gene" MAMMALIAN GENOME, vol. 4, no. 4, 1993, pages 211-219, XP002058280 see abstract see page 212 -----	1-11

☐ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

6

Date of the actual completion of the international search	Date of mailing of the international search report
13 March 1998	27/03/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Galli, I

Form PCT/ISA/210 (second sheet) (July 1992)